National Aeronautics and
Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

Multi-dataset collection research scenarios –

A Merged Atmospheric Water Data Set from the A-Train
and
A Multi-Sensor Water Vapor Climate Data Record Using Cloud Classification

Eric Fetzer

Jet Propulsion Laboratory, California Institute of Technology

CEWIS Workshop, GSFC

15 June 2010

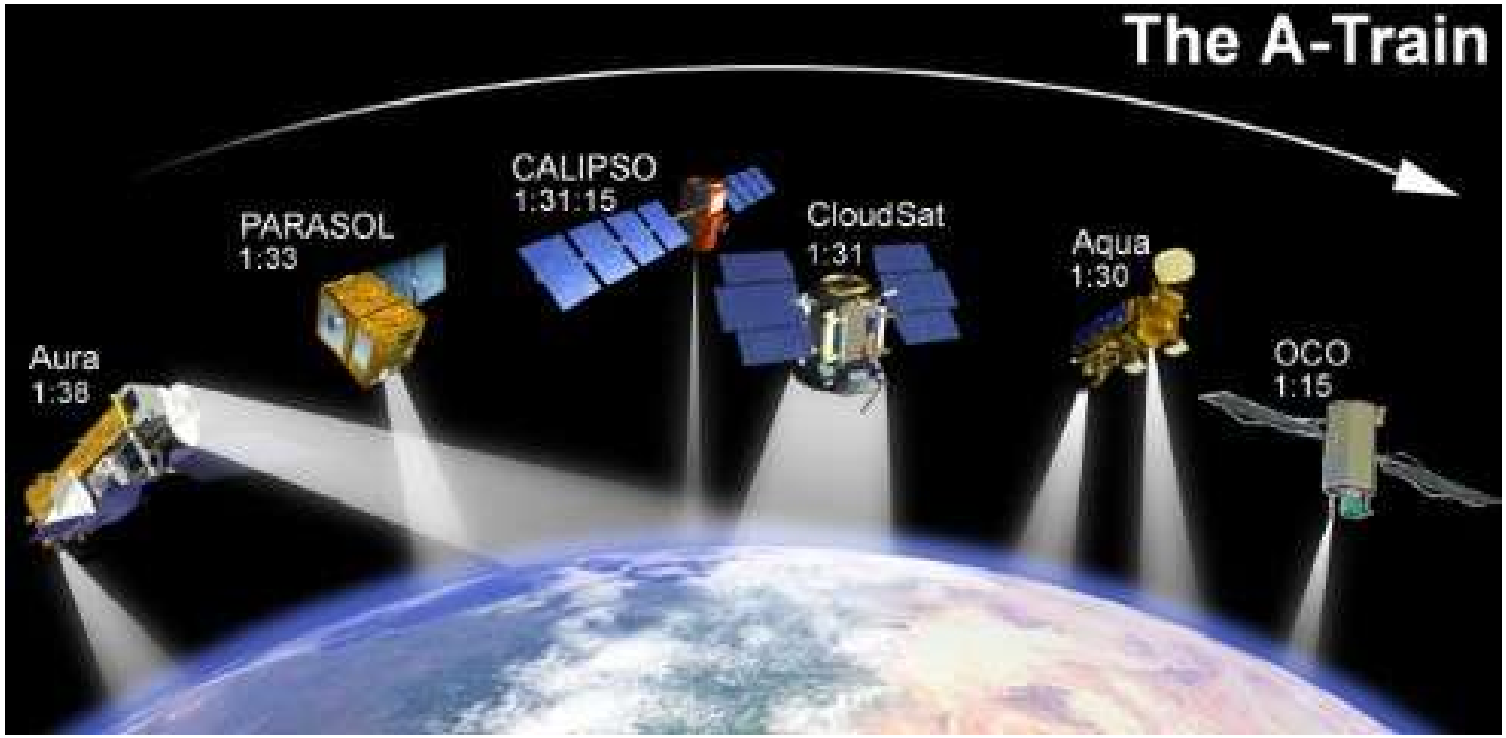
National Aeronautics and
Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

The A-Train

Multiple sensors, often identical quantities:

- **Temperature from AIRS, MLS, TES, MODIS.**
- **Water vapor from AIRS, AMSR-E, TES and MODIS.**
- **Clouds from CloudSat/CALIPSO, MODIS, AIRS and AMSR-E.**



National Aeronautics and
Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

Guiding Questions

- **What steps do you take to gather and prepare data so that you can perform multi-datasets inter-comparisons?**
- **What data related roadblocks do you encounter when bringing heterogeneous datasets together?**
- **Do you prefer to perform these services yourself, or use services provided by others?**

National Aeronautics and Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

First Guiding Question

- **What steps do you take to gather and prepare data so that you can perform multi-datasets inter-comparisons?**
  - ***Assemble A-Train Level 1 and Level 2 data sets.***
    - Since we are creating L3-type data sets.
  - ***Collocate observations from different instruments.***
  - ***Understand the many caveats, particularly with Level 2.***
    - Sensitivity, sampling, quality flags, uncertainties, etc…
  - ***Create statistical summaries.***
    - Means.
    - Clusters.
    - Conditioned means (e. g., AIRS water vapor by CloudSat classes).

National Aeronautics and
Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

Second Guiding Question

- **What data related roadblocks do you encounter when bringing heterogeneous datasets together?**
  - ***Few or no collocated Level 1 and Level 2 A-Train dataset provided by instrument teams.***
  - ***Expert knowledge of several sensors is needed for interpretation.***
  - ***Coverage is global; records are 8+ years.***
  - ***Error estimation is still in the Paleolithic.***
  - ***Reconciling and cross-linking L1 & L2 observations is THE challenge.***

National Aeronautics and
Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

Third Guiding Question

- **Do you prefer to perform these services yourself, or use services provided by others?**
  - ***We perform most of these services ourselves, guided by science insights. Services include:***
    - Winnowing through long lists of variables.
    - Matching observations.
    - Reconciling observations.
  - ***Why?***
    - The combinatorial explosion is a real possibility
      - *Several instruments X dozens to hundreds of variables = complexity.*

      $\Rightarrow$*be selective based on understanding.*
  - ***We are sharing these data sets with interested users.***

National Aeronautics and Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

Reconciling the simplest retrieved quantity: Total Precipitable Water Vapor

- **Use eight-day averaged (octad) AIRS and AMSR-E data.**
- **Examine subtropics where AIRS performs well.**
- **Examine regional five-year times series of:**
  - ***AMSR-E and AIRS total water over ocean***
  - ***Height resolved temperature and humidity from AIRS.***
- **Can we find trends anywhere?**

National Aeronautics and Space Administration

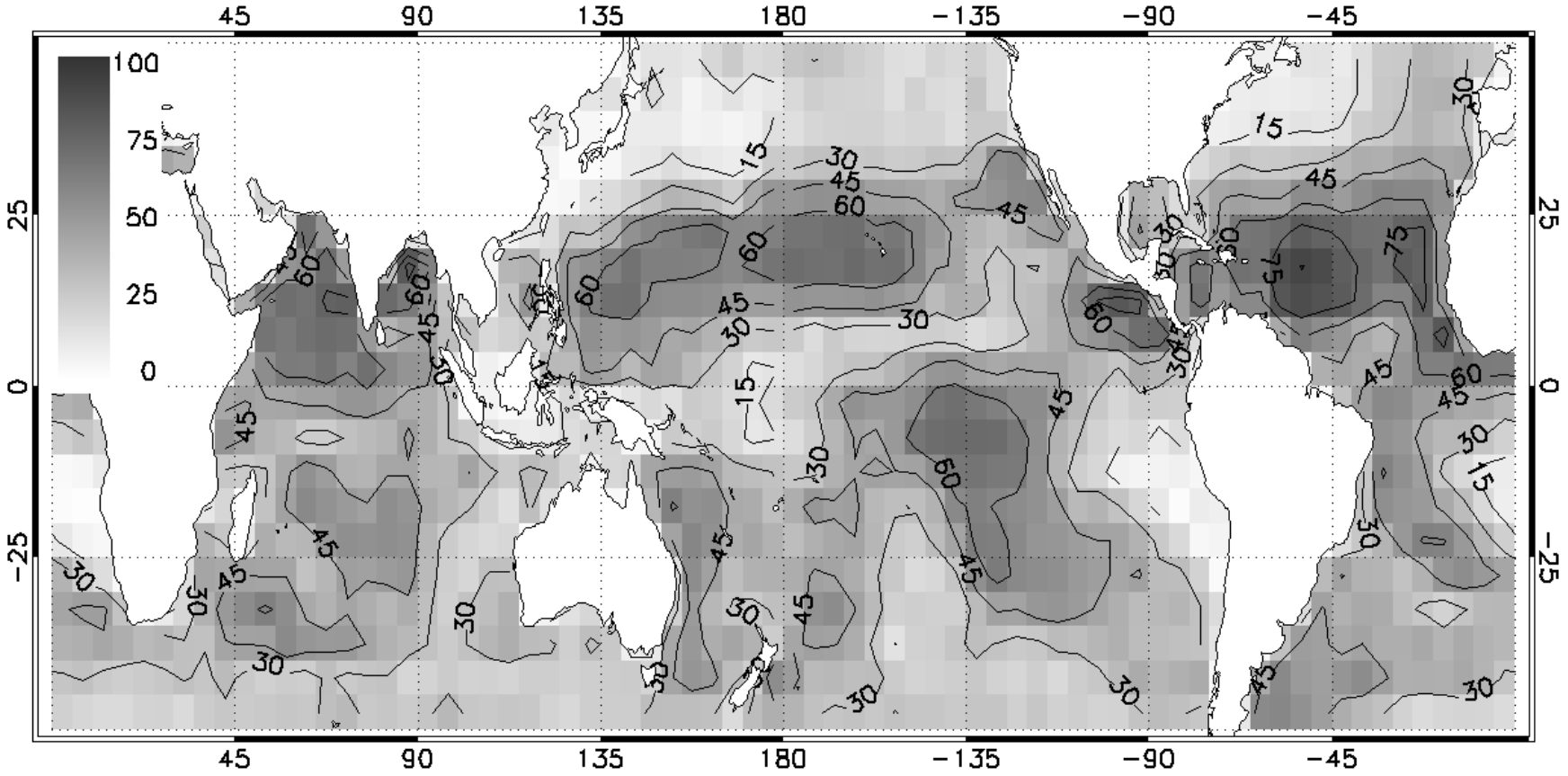
Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

AIRS retrieval yields vary with location

Fraction of 'good' retrievals (percent)

25 Dec 2002 to 15 Jan 2003

From Fetzer et al. (2006)



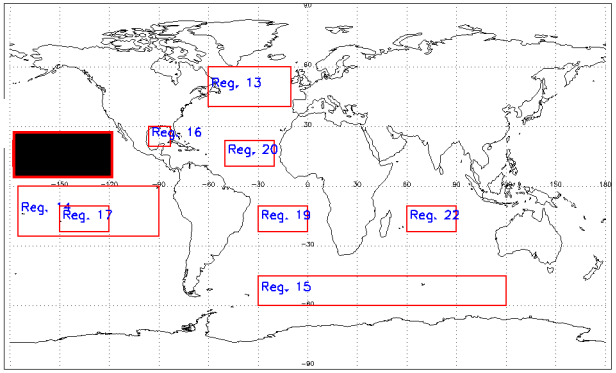
Analogous AMSR-E yields are >90% everywhere

National Aeronautics and Space Administration

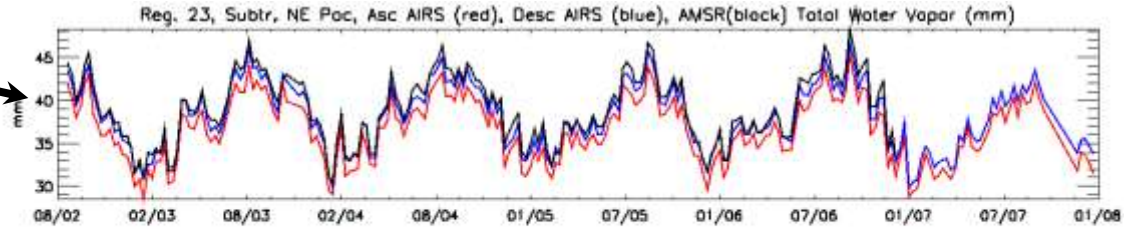
Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

Plots show 5-year times series of total water vapor from AMSR-E & AIRS

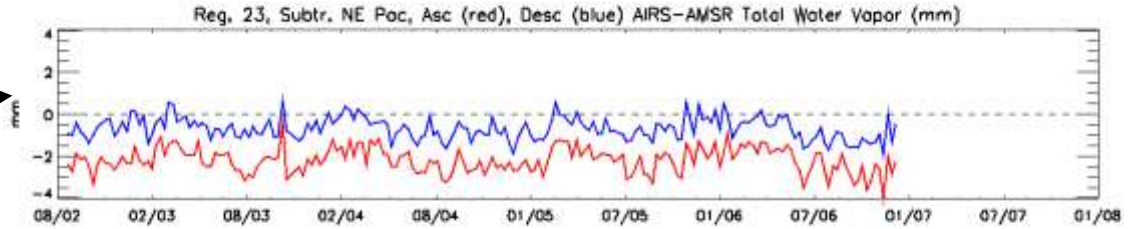
Where:



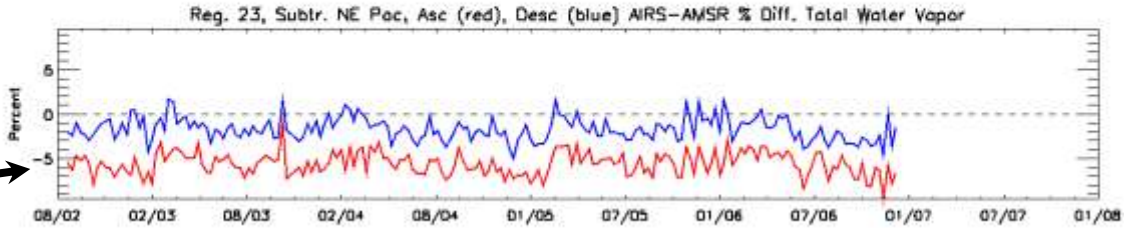
AIRS Daytime (red)
AIRS Nighttime (blue)
AMSR-E combined (black)
In mm



AIRS-AMSR-E Daytime (red)
AIRS-AMSR-E Nighttime (blue)
In mm.



AIRS-AMSR-E Daytime (red)
AIRS-AMSR-E Nighttime (blue)
In percent.

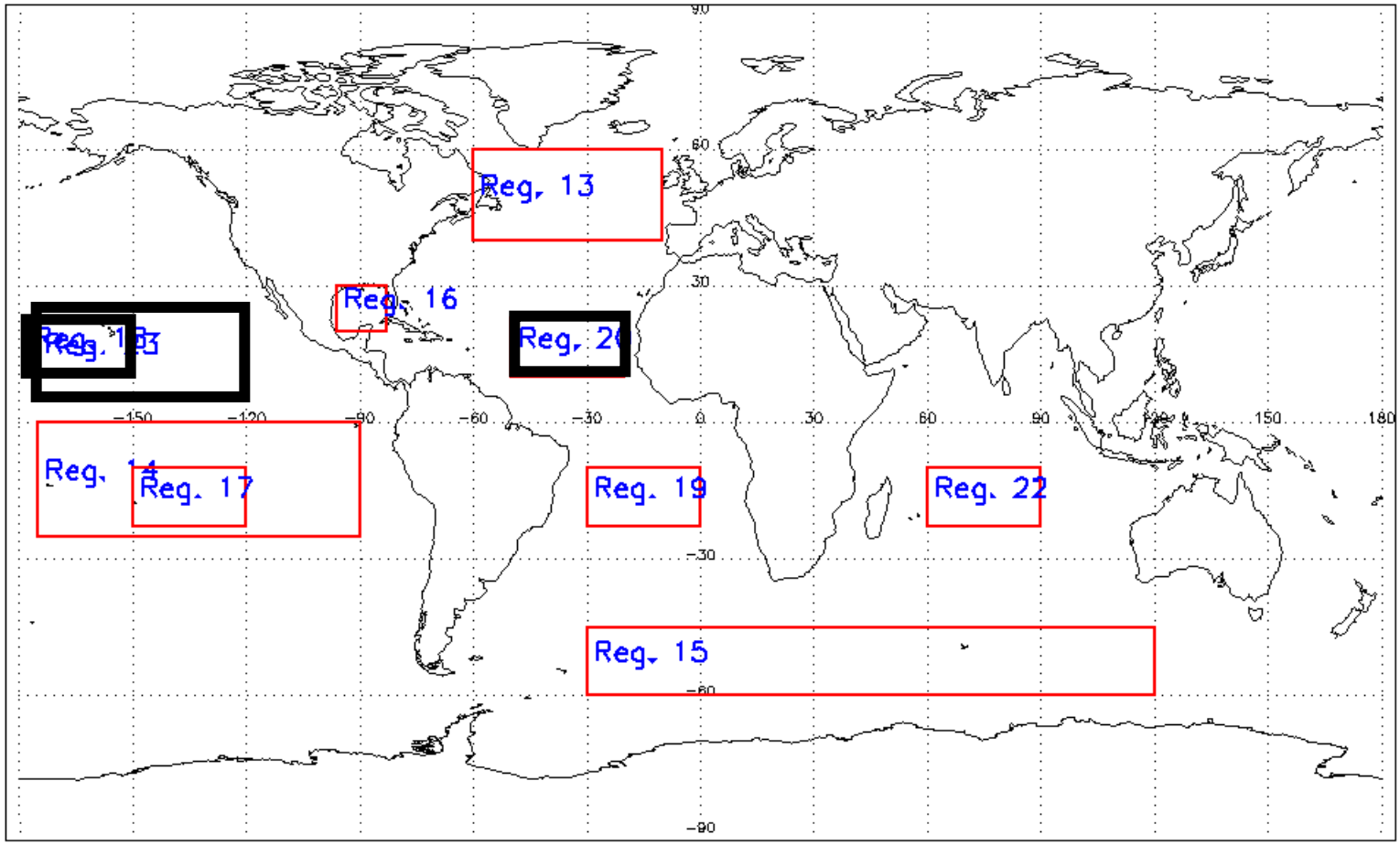


Dates: Sep 02-Mar 08

National Aeronautics and Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

Ocean Regions of Interest
Northern subtropics first

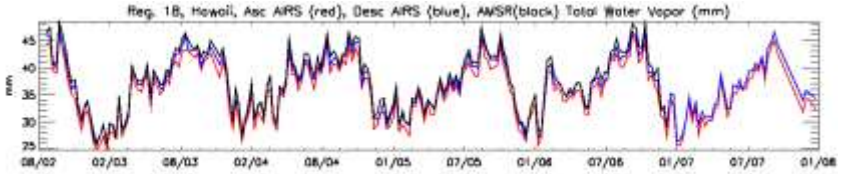


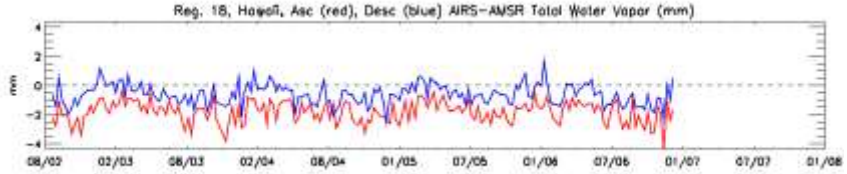
National Aeronautics and Space Administration

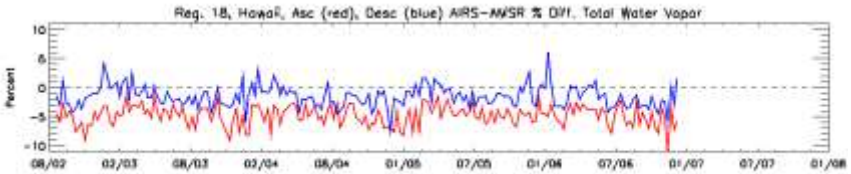
Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

Northern Subtropical Ocean

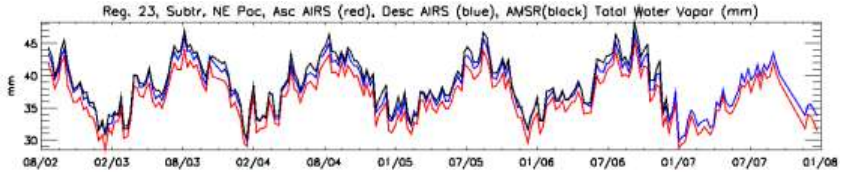
Hawaii

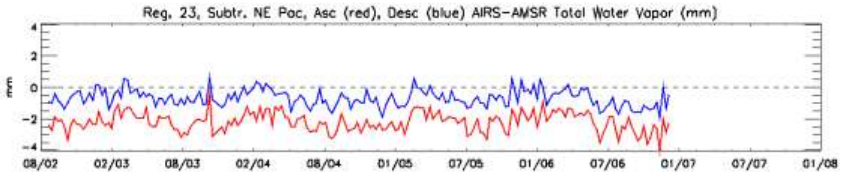


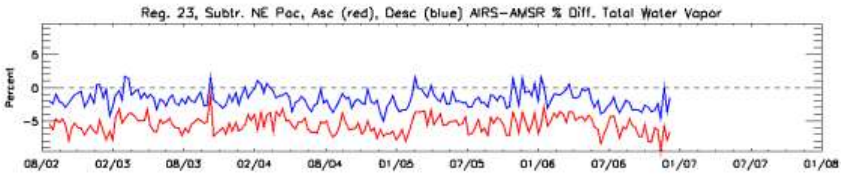


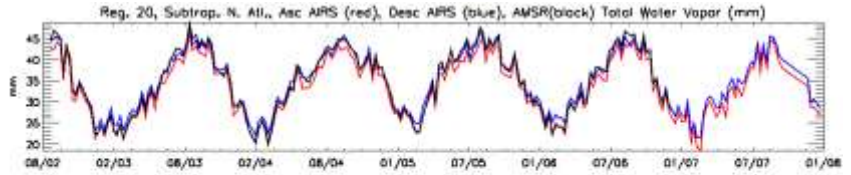


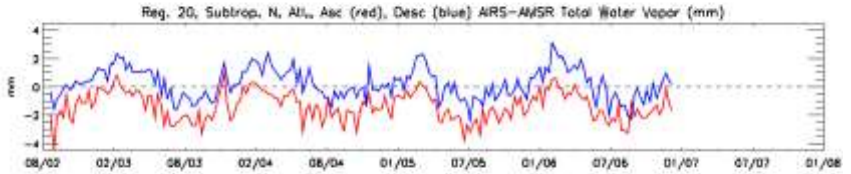
Subtropical NE Pacific (just shown)

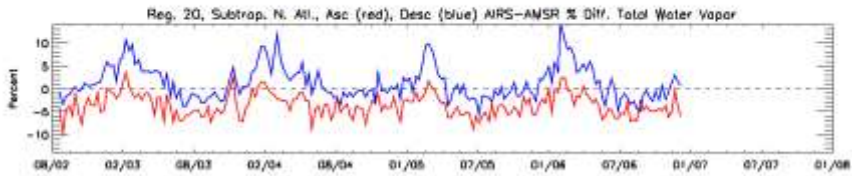










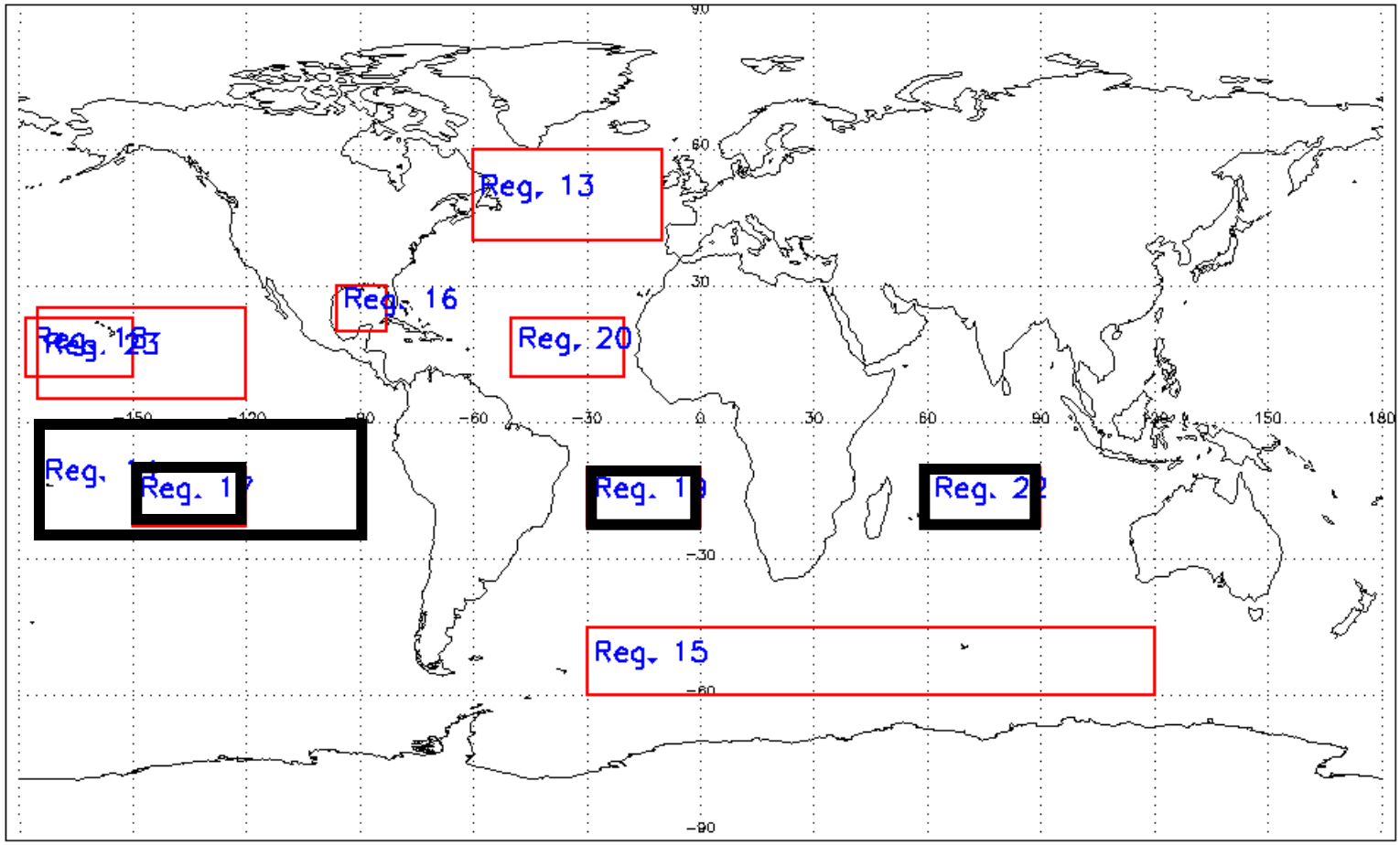


West of Africa

National Aeronautics and
Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

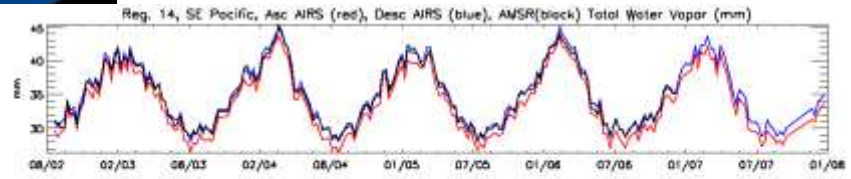
Southern Subtropical Ocean

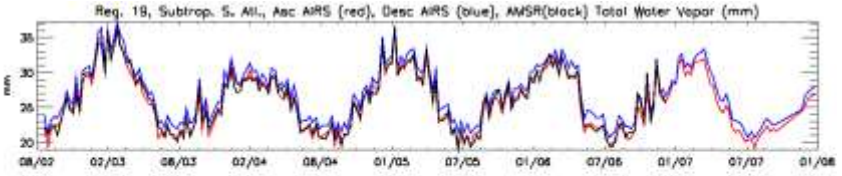


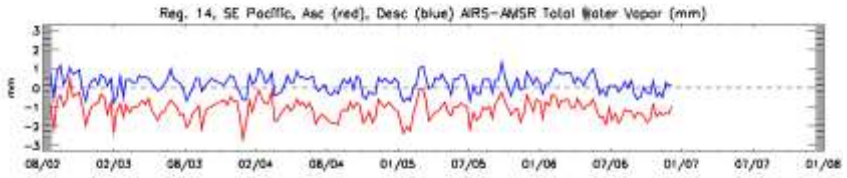
National Aeronautics and Space Administration

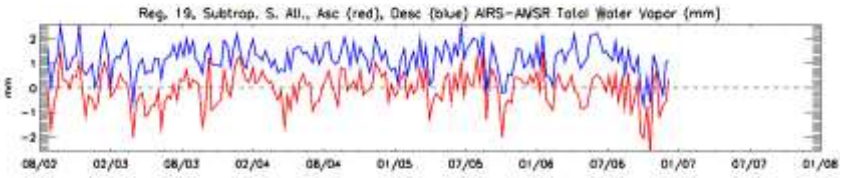
Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

Southern Subtropical Ocean

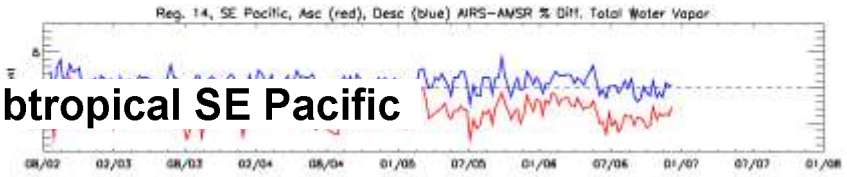


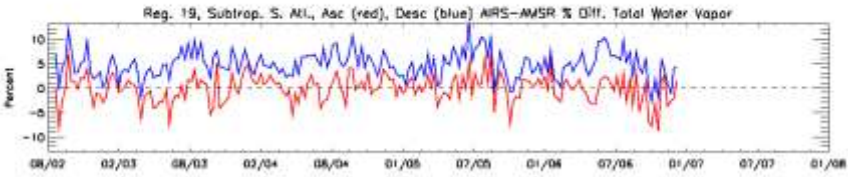




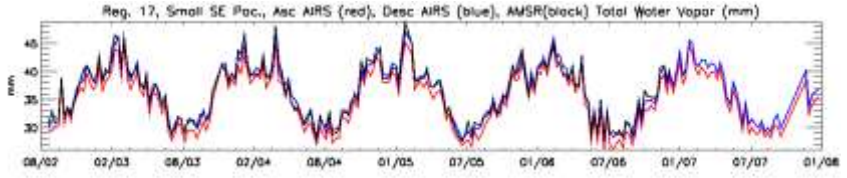


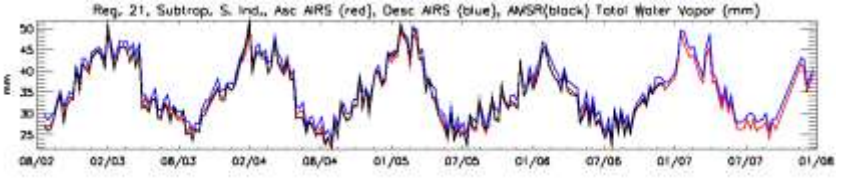
S.

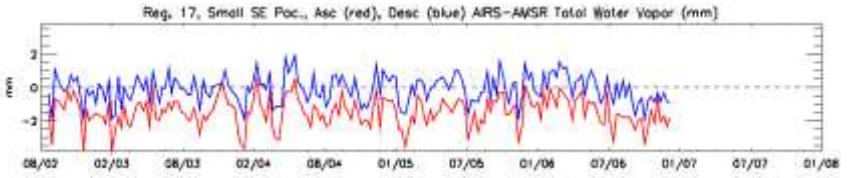


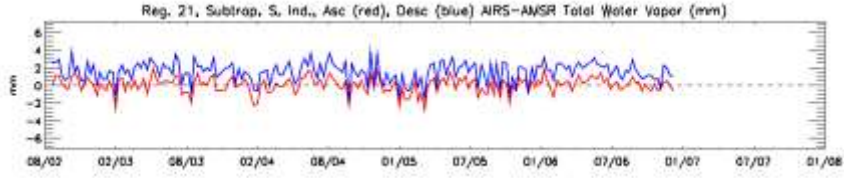


Subtropical SE Pacific

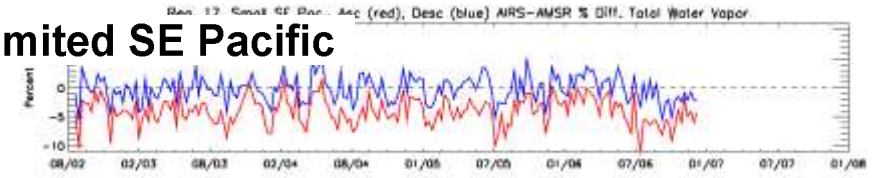


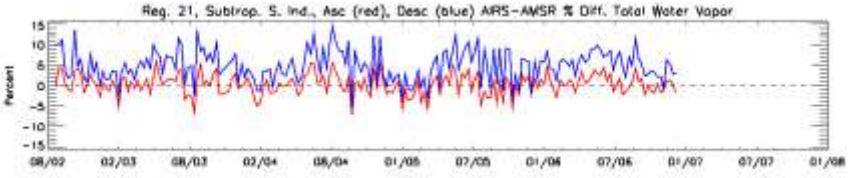






S.





Limited SE Pacific

National Aeronautics and Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

Summary of Subtropical Total Water Vapor

- **AIRS daytime is always drier than AIRS nighttime by ~3%.**
  - ***This may be an artifact of reflected shortwave IR by clouds.***
    - Testable by looking at scan angle dependency.

- **AIRS-AMSR-E difference:**
  - ***varies with location, hemisphere***
    - at most a few percent absolute.
  - ***varies with season***
    - related to changes in cloud cover (sampling effects).

- ***Conclusion*: View local trends with caution because of AIRS day-night differences and AIRS-AMSR-E biases**
  - ***Note: true AIRS-AMSR-E bias is indeterminate from this study.***

National Aeronautics and
Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

Reconciling water vapor observations is a challenge.

National Aeronautics and
Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

Reconciling cloud observations is generally harder.

National Aeronautics and Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

Cloud Fraction

The fraction of the sky that is covered by clouds; also referred to as Cloud Amount or Cloud Cover. Cloud Fraction or Cloud Cover is defined as the number of cloudy pixels divided by the total number of pixels. It also refers to the amount of sky estimated to be covered by a specific cloud type (cirrus, contrail, cirrostratus, cumulus, ...), cloud particle phase(water, ice, liquid, mixed), cloud height (low, mid, high), or by all cloud types (total cloud fraction). Cloud differs from fog only in that the latter is, by definition, close (a few meters) to the Earth's surface.

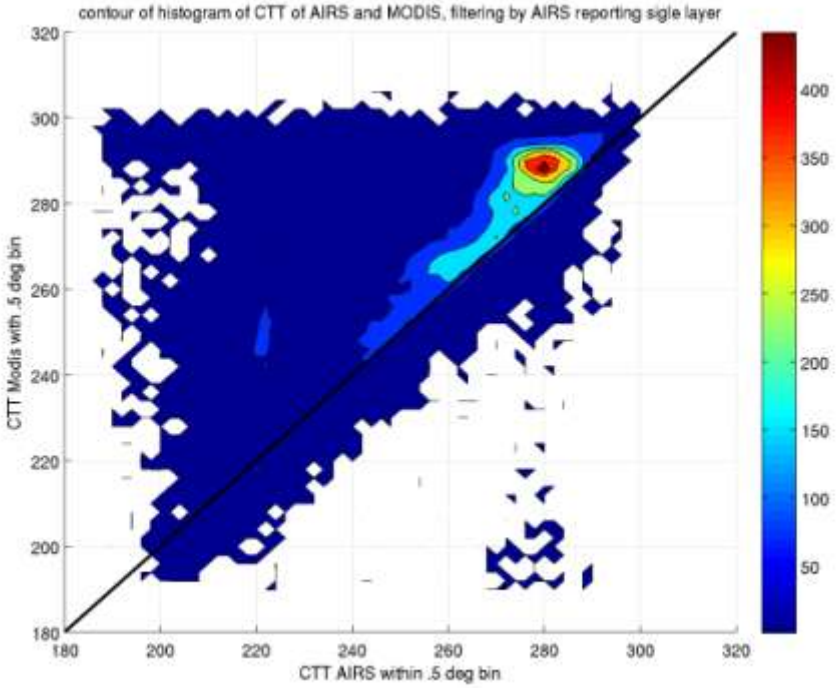
Data Set	Description	Date Range	Number of Items	Avg Size (MB)
PathA NOAA-12 (ND) Daily info	Morning orbit: 1930 LST ascending node, 0730 LST descending node	1991-07-04 to 1994-07-01	1816	9.357
MAT3CPRAD.5.2.0 info	tavg3_3d_rad_Cp: MERRA Assimilation 3D Incremental Analysis Update (IAU) Diagnostic, Upper-air diagnostics from radiation, Time average 3-hourly on 1.25x1.25 grid **Available Services:** • Download via HTTP	1979-01-01 to 2009-12-31	11323	85.776
OMCLDO2_CPR.003 info	OMI/Aura Level 2 Cloud Pressure and Fraction (O2-O2 Absorption) along CloudSat track Version 3	2006-06-01 to 2010-03-31	19586	1.932
PathA NOAA-10 (NG) Daily info	Morning orbit: 1930 LST ascending node, 0730 LST descending node	1986-11-25 to 1991-09-01	3480	9.349
PathA NOAA-11 (NH) Five_day_Mean info	Afternoon orbit: 1340 LST ascending node, 0140 LST descending node	1988-10-12 to 1994-06-30	379	20.587
PathB NOAA-10 (NG) Daily		1986-12-31 to 1991-09-01	2555	2.72

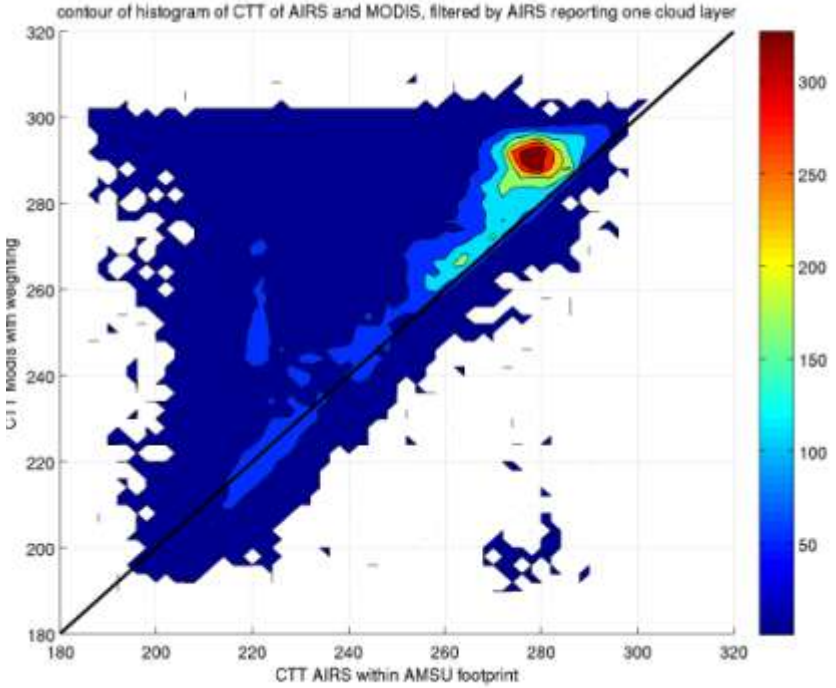
Page 2 of 10

National Aeronautics and Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

NEWS Merged A-Train Data





The merged 0.5 degree data set (left), is very representative of the behavior of the original data (right), but is many times smaller. The merged 0.5 degree data set averages observations only within seconds of each other, and at most 0.5 degree apart. The merged data set preserves much of the variability associated with change in time and space, but is much less cumbersome. The merged data set has also been quality filtered to eliminate observations considered unusable by the different instrument teams. This reduces user need for familiarity with multiple Level 2 data sets.

National Aeronautics and Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

Providing Web-based services based on the NEWS data sets

"Web Services-enabled Tool for Distributed Custom Level 2 Data Subsetting and Level 3 Data Summarization"

RESPONDING TO A NASA RESEARCH ANNOUNCEMENT (NRA) and Appendix A.22

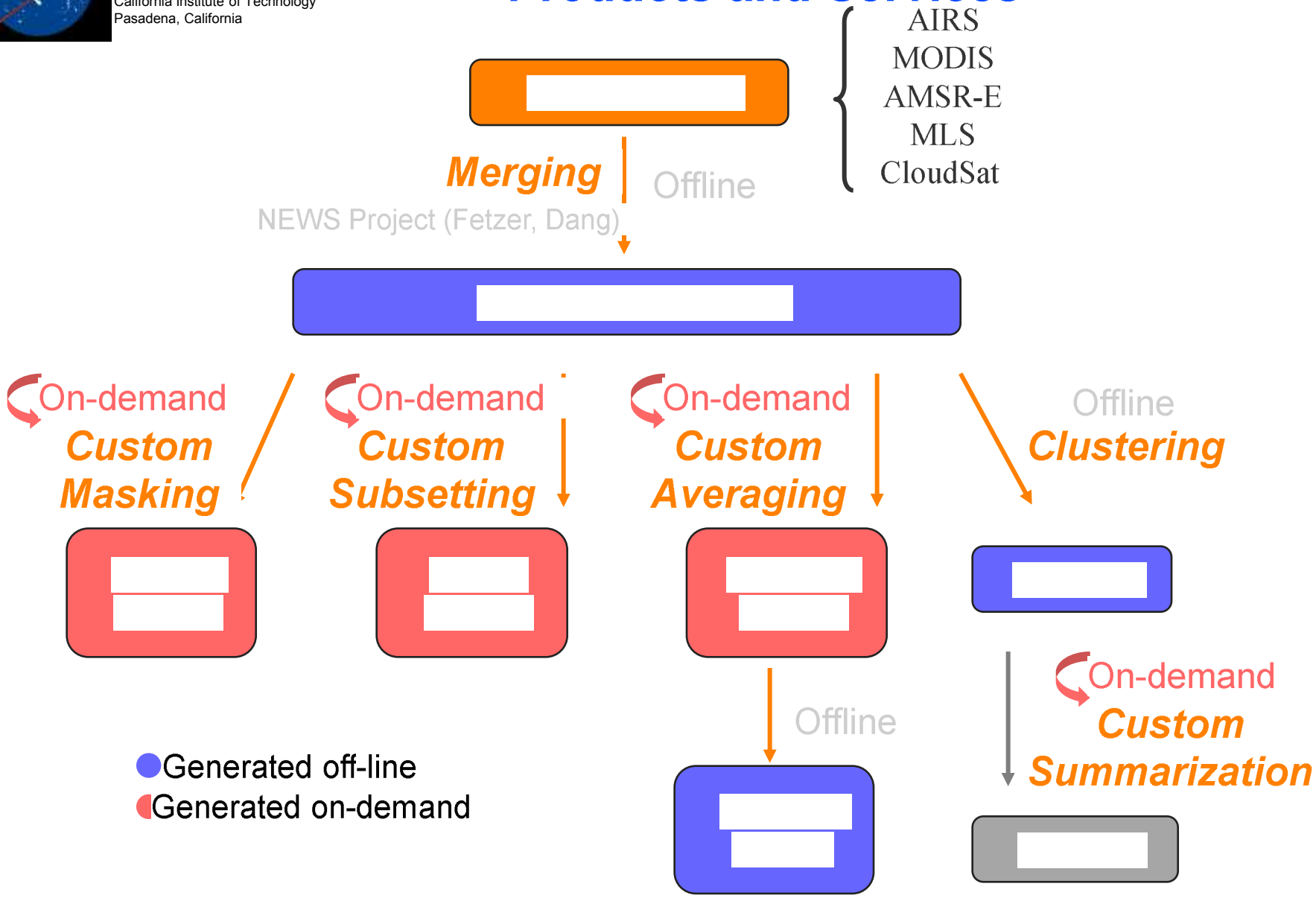
P. I. Hook Hua, JPL

Two-year project just ending…

National Aeronautics and
Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

Products and Services

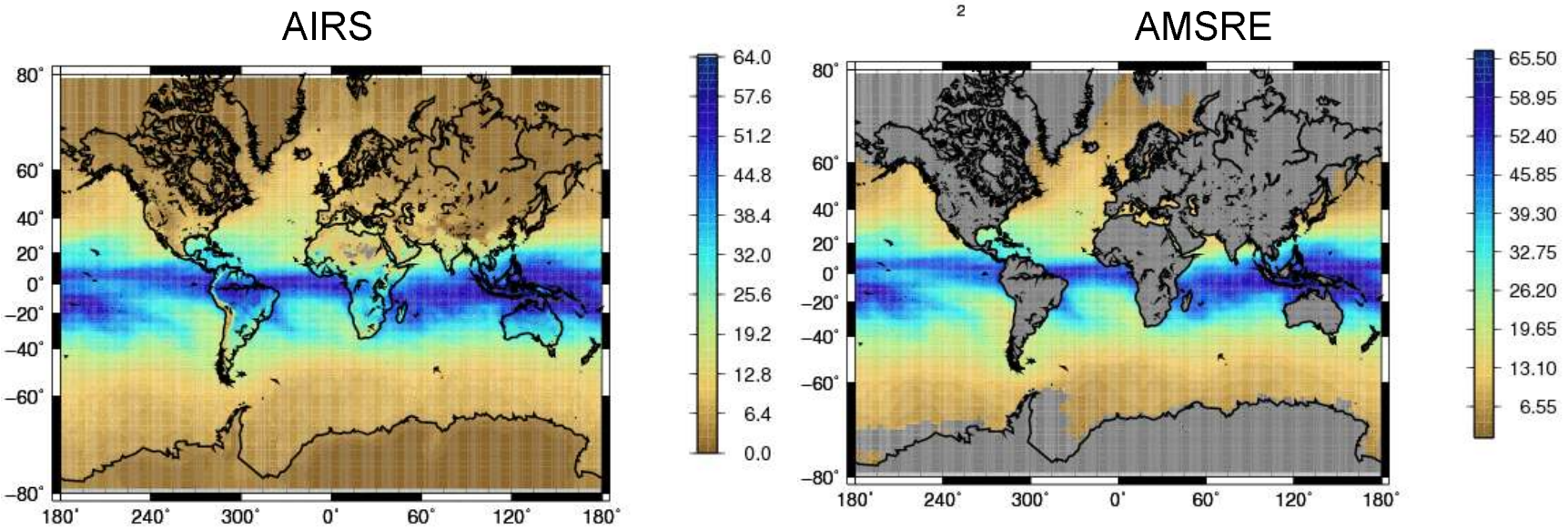


National Aeronautics and Space Administration
Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

On-Demand Averaging Service

- **Average data that meet a temporal (time range) condition and parameter choices that a user specified on a grid with a customizable resolution.**
- **A user can specify the grid resolution with any multiple of 0.5 degrees.**

Averaged Total Water Vapor for 2004 February on 1x1 degree grid



National Aeronautics and Space Administration
Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

On-Demand Subsetting Service

- **Collect data that meet temporal (time range) and spatial (lat-lon bounding box) conditions and parameter choices that a user specified.**

Instrument	Parameters
AIRS	Air temperature profile in atmosphere layer Humidity mixing ratio Atmosphere water vapor content Cloud area fraction profile in atmosphere layer Cloud top air temperature
AMSRE	Atmosphere water vapor Atmosphere cloud liquid water content Sea surface temperature Rainfall rate
MLS	Humidity mixing ratio Atmosphere cloud ice content Cloud ice content of atmosphere layer
MODIS	Atmosphere cloud condensed water content Atmosphere cloud liquid water content Atmosphere cloud undetermined condensed water content Cloud effective radius Cloud optical thickness Cloud area fraction in atmosphere layer Air temperature at cloud top

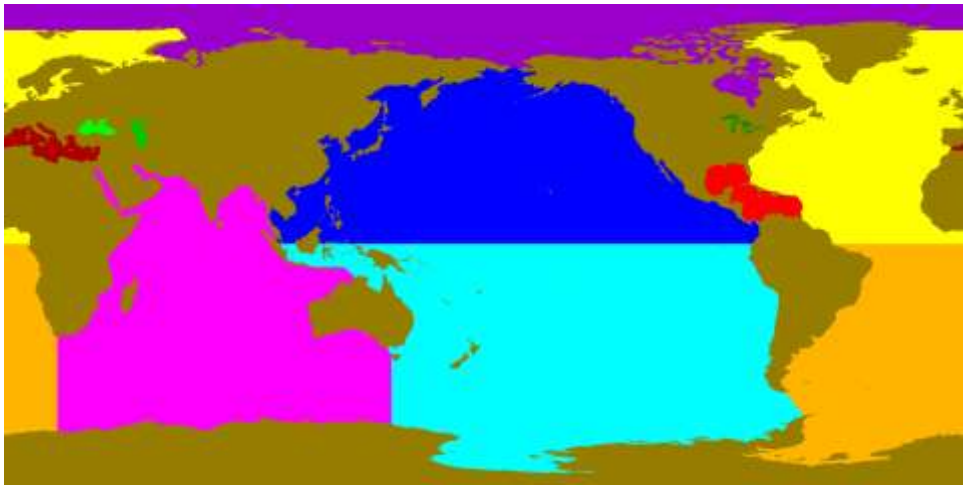
National Aeronautics and Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

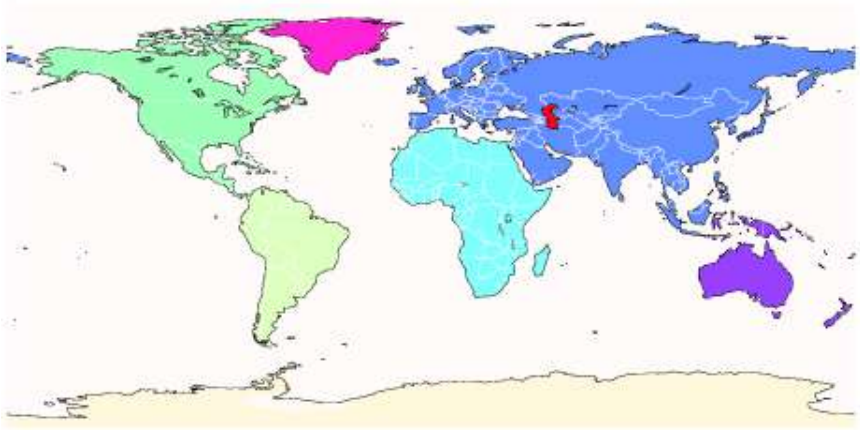
On-Demand Masking Service

- **Collect data that meet temporal (time range) and spatial (geographical regions) conditions and parameter choices that a user specified.**
- **The geographical regions that the user can choose from are water, land, Antarctica, South America, North America, Africa, Eurasia, Australia, Arctic, Caribbean, Mediterranean, Black Sea, Caspian Sea, Great Lakes, North Pacific, South Pacific, Indian, North Atlantic, and South Atlantic.**

Ocean basins



Continents

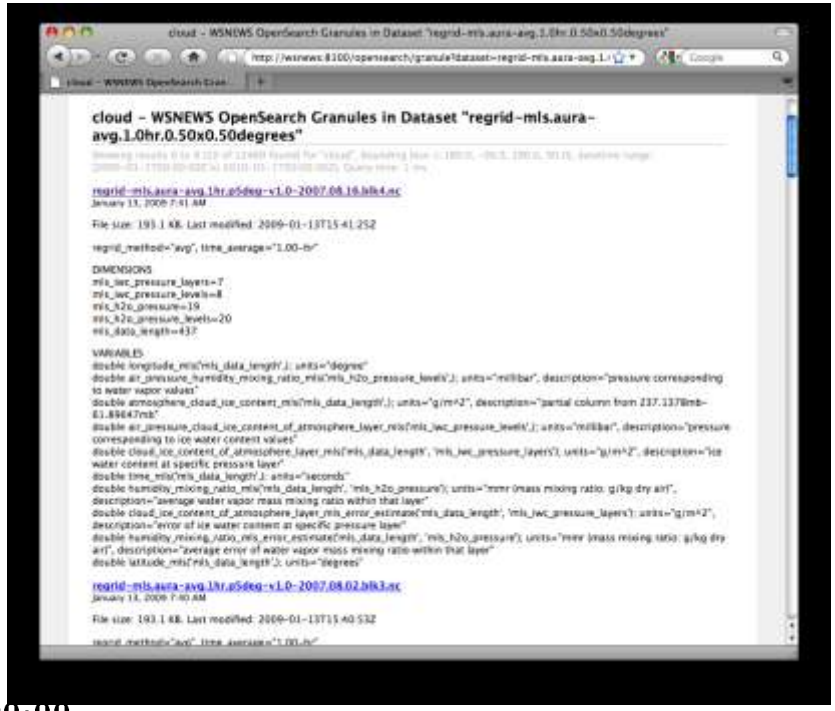


National Aeronautics and Space Administration
Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

OpenSearch Data Access Service

- **Free-text + Space + Time search of merged A-Train data**
- **Currently about 34K granules available**
  - ***Level 2 Regrid***
  - ***Level 3 Averages***
  - ***Level 3Q Summaries***
- **Dataset search**
- **Granule search**
- **Direct downloads**

Search Results as News Feed

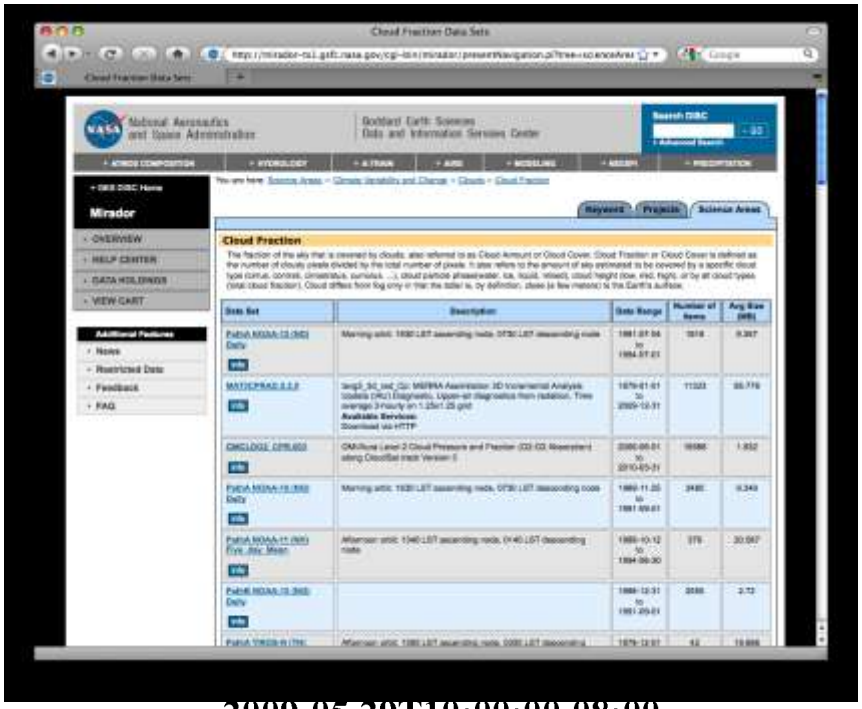


National Aeronautics and Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

Integration with Mirador

- **Real-time search of NEWS data at JPL from Mirador at GSFC via OpenSearch service**
- **Allows drilling down to parameter**
- **Pulls in external data sources like WSNEWS**



2009-05-29T10:00:00-08:00

National Aeronautics and
Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

Data Provenance also Important

Tracking Production Legacy of Multi-Sensor Merged Climate Data Records

PROPOSAL FOR
ACCESS (ROSES 09 A.34)
RESPONDING TO A
NASA RESEARCH ANNOUNCEMENT (NRA)